

# Tighter Bounds on the Inefficiency Ratio of Stable Equilibria in Load Balancing Games

Akaki Mamageishvili and Paolo Penna

Department of Computer Science, ETH Zurich, Switzerland

November, 2015

## Abstract

In this paper we study the inefficiency ratio of stable equilibria in load balancing games introduced by Asadpour and Saberi [3]. We prove tighter lower and upper bounds of  $7/6$  and  $4/3$ , respectively. This improves over the best known bounds in problem (19/18 and  $3/2$ , respectively). Equivalently, the results apply to the question of how well the optimum for the  $L_2$ -norm can approximate the  $L_\infty$ -norm (makespan) in identical machines scheduling.

## 1 Introduction

Load balancing problems are actively studied in the context of games. Transition from a centralized optimization problem to a decentralized environment is immediate, by assuming that the jobs are owned by selfish but rational players. These games are prototypical of resource allocation problems in which users (players) do not act altruistically, therefore leading the system to suboptimal configurations. Naturally, one can consider the *social optimum* as the allocation minimizing the *makespan*, that is, the maximum load over all machines (a classical measure of efficiency). In contrast, in the game-theoretic setting, the players will optimize for their own job only, typically reaching what is called a *Nash equilibrium*, that is, an allocation such that no player has an interest in moving to another machine.<sup>1</sup> The study of how inefficient these Nash equilibria can be is a central problem in algorithmic game theory, as it measures how much selfishness can impede optimization.

Asadpour and Saberi [3] introduced and studied the *inefficiency ratio of stable equilibria (IRSE)* in several games, including load balancing ones (see Section 2 for formal definitions). This notion quantifies the efficiency loss in games when players play certain *noisy* best-response dynamics (see Section 1.1 for a detailed discussion). For load balancing games, the IRSE has another very simple and natural interpretation (which is also of independent interest and studied earlier):

---

<sup>1</sup>In this work we consider *pure* Nash equilibria, that is, configurations in which each player chooses one strategy and unilateral deviations are not beneficial. In general, games may also possess mixed Nash equilibria in which players choose strategies according to a probability distribution.

*Are the allocations minimizing the  $L_2$ -norm (sum of the squares of the machine loads) also sufficiently good for minimizing the  $L_\infty$ -norm (makespan)?*

Intuitively, the IRSE on load balancing games is equal to some value  $c$  if *every* allocation minimizing the  $L_2$ -norm is automatically a  $c$ -approximation for the  $L_\infty$ -norm (i.e., the makespan of this allocation is at most  $c$ -times the optimal makespan). An exact bound on the IRSE is not known (as opposed to other measures related to the inefficiency of equilibria – see next section). Asadpour and Saberi [3] proved an upper bound of  $\frac{3}{2}$  on IRSE and observed that an example in [1] implies a lower bound on IRSE of  $\frac{19}{18}$ . In this work we improve both bounds: In Section 3 we show an improved lower bound of  $7/6$ , while the upper bound of  $4/3$  is proved in Section 4.

## 1.1 Significance of the results and related work

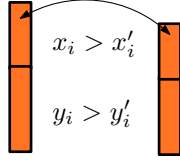
The inefficiency of Nash equilibria is often measured through two classical notions: The *price of anarchy* ( $PoA$ ) introduced in Koutsoupias and Papadimitriou [8] for load balancing games on related machines, and the *price of stability* ( $PoS$ ) introduced in Anshelevich et al. [2], which compare respectively the *worst* and the *best* Nash equilibrium to the social optimum. In some cases, these notions can be considered too extreme as they may include “unrealistic” equilibria.

The IRSE [3] studies the quality of equilibria selected by certain *noisy* best-response dynamics [4]. Intuitively, these dynamics will most likely rest on pure Nash equilibria *minimizing the potential* of the game, and the IRSE can be seen as the price of anarchy restricted to these selected equilibria (in fact, IRSE is also known in the literature under the name of *potential optimal PoA* by Kawase and Makino [7], who also considered the analogous *potential optimal PoS*). In load balancing games (and several others) we have  $PoS < IRSE < PoA$ , which in a sense tells that the  $PoS$  and the  $PoA$  are either too pessimistic or too optimistic. Specifically, on  $m$  machines  $PoA = (2 - \frac{1}{m+1})$  [5, 9],  $PoS = 1$  [9], while  $IRSE$  is between  $19/18$  and  $3/2$  [3]. The latter bounds are strengthened in the present paper to  $7/6 \leq IRSE \leq 4/3$ . This means that players can easily compute a  $4/3$  approximation of the optimum via these simple dynamics, but also that optimal (exact) solutions are unlikely to be chosen either, in some instances.

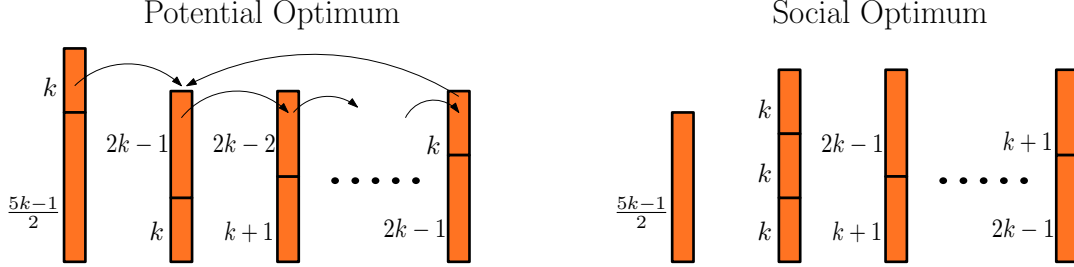
The upper bound  $4/3$  also suggests an intriguing comparison with the study of *sequential PoA* by Hassin and Yovel [6] for these games: There players are far from myopic and the equilibrium is meant on an extensive form games in which players are able to reason about future moves of following players. Our upper bound says that myopic best-response are not worse than in their dynamics where players are capable of more sophisticated reasoning.

## 2 Preliminaries

In load balancing we have  $n$  jobs of weight  $w_1, \dots, w_n$  that we want to put on  $m$  identical machines (each job is allocated to one machine). The job allocation determines the load  $l_j$  of each machine  $j$ , that is the sum of the jobs weights that are allocated to this machine. The goal is to find an allocation that has the lowest possible *makespan*, that is, the maximum load over all machines.



**Figure 1:** When a swap of bundles of jobs reduces the potential.



**Figure 2:** Lower bound  $7/6$  on IRSE.

In *load balancing games* each job is a *player* who can choose any of the  $m$  possible machines. The cost for player  $i$  is simply the load of the machine chosen by this player, and naturally each player aims at minimizing *her own cost*. The strategies of all players specify a job allocation (in the game theoretic terminology this is the strategy profile).

An allocation minimizing the makespan is called a *social optimum*, and its makespan is called a *social optimum makespan*. The *potential* associated to an allocation is the sum of the *squares* of the corresponding machine loads,  $l_1^2 + \dots + l_m^2$  where  $l_j$  is the load of machine  $j$  at this allocation. An allocation minimizing the potential function is called a *potential optimum*.

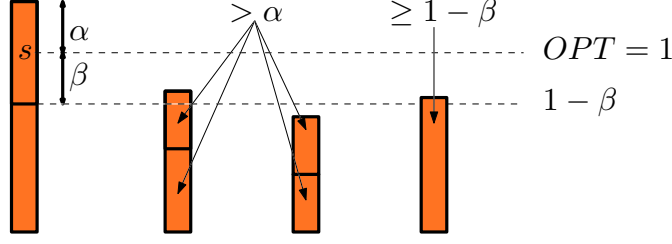
It is well known that load balancing games are weighted potential games with the above potential function. This means that all pure Nash equilibria, allocations where no player can unilaterally improve moving to another machine, are actually ‘local optimal’ for the potential (a single job move cannot reduce the potential). Asadpour and Saberi [3] introduced and studied the *inefficiency ratio of stable equilibria (IRSE)*, which is the ratio between the *worst* makespan of a potential optimal allocation divided by the social optimum makespan.

Potential optimum allocations satisfy the following condition (see Figure 1). Split the total load of each machine into two bundles of jobs, that is,  $l_j = x_j + y_j$  where  $x_j$  is the sum of the weights of a (possibly empty) subset of jobs allocated to  $j$ . If two machines  $i$  and  $j$  satisfy  $x_i < x_j$ , then  $y_i > y_j$  for otherwise swapping  $x_i$  with  $x_j$  reduces the potential. Pure Nash equilibria satisfy the weaker condition that a single job in machine  $i$  does not improve if moving to another machine  $j$ , that is,  $l_i - w_i \leq l_j$ .

### 3 Improved lower bound

In this section we strengthen the  $19/18$  lower bound on IRSE in [3, 1].

**Theorem 3.1.** *IRSE in load balancing games is at least  $7/6$ .*



**Figure 3:** Proof of upper bound – Lemma 4.2.

*Proof.* Consider the instance in Figure 2 where  $m = k + 1 \geq 3$ ,  $n = 2k + 2$  and the weights of jobs are  $k, k, k, k + 1, k + 1, \dots, 2k - 1, 2k - 1, (5k - 1)/2$ . We prove that the left allocation of jobs on the figure minimizes the potential function, while the one on the right has optimal makespan, thus implying

$$IRSE \geq \frac{(7k - 1)/2}{3k} = \frac{7m - 8}{6m - 6}$$

which tends to  $7/6$  as  $m$  goes to infinity. First note that the potential in both allocations are equal. Consider any job allocation and without loss of generality assume that the job with the largest weight is on machine 1. If we fix a load on machine 1, then any job allocation which balances the load on the other machines minimizes the potential (among all allocations with this fixed load on machine 1). Therefore, if the largest job is alone on machine 1 then the potential is minimized in the social optimum case, while if it is located on machine one together with job of smallest weight then the potential is optimized when the job allocation is like on the left side of Figure 2. If the load on machine 1 is strictly larger than  $5(k - 1)/2 + k$  then the potential is not a potential optimal allocation. We conclude that both allocations are potential optimal, in particular the one on the left which gives the lower bound on IRSE.  $\square$

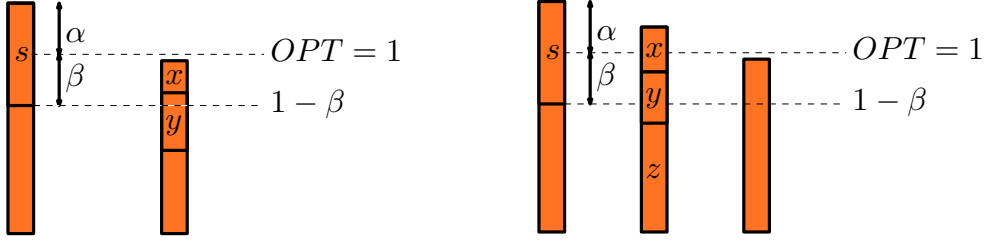
## 4 Improved upper bound

Without loss of generality assume that a makespan in social optimum is equal to 1 and loads of machines in the potential optimizer job allocation are sorted in a non-increasing order  $l_1 \geq l_2 \geq \dots \geq l_m$ .

**Theorem 4.1.** *IRSE in load balancing games is at most  $4/3$ .*

*Proof.* We have to prove that  $l_1 \leq 4/3$ . Assume that it does not hold, that is  $l_1 = 1 + \alpha > 4/3$  implying  $\alpha > 1/3$ . Machine 1 contains at least two jobs, otherwise the optimum makespan would not be 1. The smallest job on machine 1, which we call  $s$ , is strictly larger than  $\alpha$ , otherwise since  $l_m < 1$  we can decrease the potential by moving  $s$  to the last machine. Assume that the weight of  $s$  is  $\alpha + \beta$  where  $\beta > 0$  and since  $\alpha + \beta \leq l_1 - \alpha - \beta = 1 - \beta$  it implies  $\beta < \alpha$ . The proof of the theorem is based on the following property of potential optimal allocations (whose proof is postponed).

**Lemma 4.2.** *For any  $\alpha > 1/3$ , in a potential optimal allocation, every machine must have one job of weight at least  $1 - \beta$  or two jobs of weight strictly bigger than  $\alpha$ .*



**Figure 4:** The two cases to prove Lemma 4.2.

Note that for  $\alpha > 1/3$  this contradicts the fact that the optimum makespan is equal 1 since any job from machine 1 has weight larger than  $\alpha$  and together with these jobs cannot be allocated on  $m - 1$  machines without exceeding 1 (clearly the jobs in machine 1 cannot be together in the optimum makespan, no three jobs of weight  $\alpha$  can fit together, nor a job of  $1 - \beta$  with a job of  $\alpha$ ).  $\square$

#### 4.1 Proof of Lemma 4.2

We next prove Lemma 4.2. First observe that  $l_i \geq 1 - \beta$  for otherwise we can move job  $s$  from machine 1 to machine  $i$  and decrease the potential. Sort the jobs on machine  $i$  in increasing weight, and consider a bundle of small jobs such that the load is still above or equal to  $1 - \beta$  after removing them from the machine (this bundle can be empty). Let  $x$  be the overall weight of this bundle, and  $y$  be the size of the next job whose removal reduces the load below  $1 - \beta$ . That is

$$l_i - x \geq 1 - \beta \quad \text{and} \quad l_i - x - y < 1 - \beta. \quad (1)$$

Also note that

$$x + y \geq \alpha + \beta \quad (2)$$

otherwise we can exchange job  $s$  on machine 1 with job bundle of weight  $x + y$  on machine  $i$  and strictly decrease the potential. If  $y$  was the only job after removing  $x$ , then by definition  $y \geq 1 - \beta$  and the lemma holds. Otherwise, we show that  $y \geq \alpha$ . We next distinguish two cases depending on the load  $l_i$  of machine  $i$  (see Figure 4):

$(l_i < 1)$ . From the first inequality in (1) we get  $x < \beta$ , and then by (2)  $y > \alpha$ .

$(l_i \geq 1)$ . If  $x < \beta$  then  $y > \alpha > 1/3$  using the same argument of the previous case. Assume therefore  $x \geq \beta$  and  $y \leq 1/3 < \alpha$ . Let  $z$  be the load of machine  $i$  after removing the bundle of size  $x + y$ , that is,  $l_i = x + y + z$ . We rearrange the jobs in the potential optimum and show that this will further reduce the potential, thus a contradiction. The job rearrangement goes as follows:

1. Job  $s$  moves from machine 1 to machine  $i$ ;
2. The smallest between  $x$  and  $y$  moves to machine  $m$ , while the biggest moves to machine 1.

Consider first the case  $x \leq y$ . Then the difference between potential of the original allocation and potential of the new one is

$$\begin{aligned} (1 + \alpha)^2 + (x + y + z)^2 + l_m^2 - (1 - \beta + y)^2 - (z + \alpha + \beta)^2 - (l_m + x)^2 = \\ 2\left(\alpha + xy + yz + xz + \beta + \beta y - (\beta^2 + y + \alpha z + \alpha\beta + z\beta + l_m x)\right) > \\ 2\left(\alpha + xy + yz + xz + \beta + \beta y - (\beta^2 + y + \alpha z + \alpha\beta + z\beta + x)\right) \end{aligned} \quad (3)$$

where the inequality is due to  $l_m < 1$ . Next we prove that this quantity is positive and obtain the desired contradiction. Note that  $y + z \leq l_m$  otherwise moving the bundle of weight  $x$  to machine  $m$  reduces the potential. Therefore (3) is linear in  $x$  with slope  $y + z - 1 \leq 0$  and since  $x \leq y$ , we can bound it from below by

$$2\left(\alpha + y^2 + 2yz + \beta + \beta y - (\beta^2 + y + \alpha z + \alpha\beta + z\beta + y)\right) \geq$$

since this quantity is linear in  $z$  with slope  $2y - (\alpha + \beta) \geq x + y - (\alpha + \beta) \geq 0$  and  $z \geq 1 - \beta - x \geq 1 - \beta - \alpha$

$$\begin{aligned} 2\left(\alpha + 2y - \beta y - y^2 + \beta - (\beta^2 + y + \alpha - \alpha y + \beta - \beta^2 - \beta y + y)\right) = \\ 2\left(y - \beta y - y^2 + \beta - (-\alpha y + \beta - \beta y + y)\right) = \\ 2\left(\alpha y - y^2\right) \end{aligned}$$

which is strictly positive since  $y < \alpha$ .

The case  $y < x$  is similar to the previous one. Specifically, now the potential difference is strictly bigger than (3) but with  $x$  and  $y$  exchanged,

$$2\left(\alpha + xy + yz + xz + \beta + \beta x - (\beta^2 + x + \alpha z + \alpha\beta + z\beta + y)\right) \quad (4)$$

and this quantity is linear in  $y$  with non-positive slope  $x + z - 1 \leq 0$ , since otherwise moving bundle  $y$  to machine  $m$  would reduce the potential. For the case  $x < \alpha$  the rest of the proof is identical to the previous case with  $x$  and  $y$  exchanged. For  $x \geq \alpha$  we use  $y < \alpha$  and obtain a lower bound for (4) by replacing  $y$  with  $\alpha$ , and then  $z$  by  $1 - \beta - \alpha$ :

$$\begin{aligned} 2\left(\alpha + \alpha x + xz + \beta + \beta x - (\beta^2 + x + \alpha\beta + z\beta + \alpha)\right) = \\ 2\left(\alpha + \alpha x + x - x(\alpha + \beta) + \beta + \beta x - (\beta^2 + x + \alpha\beta + \beta - \beta(\alpha + \beta) + \alpha)\right) = 0 \end{aligned}$$

where in the second step we use  $x \geq \beta$  since  $\beta < \alpha$  as observed before this lemma.

This completes the proof of Lemma 4.2.

## 5 Conclusions

In this paper we improve lower and upper bounds on the inefficiency ratio of stable equilibria in load balancing games, while still leaving a gap. We believe that our lower bound example can not be improved. In order to prove this claim one has to consider global properties of potential optimizer, while in our proof only local properties have been used.

**Acknowledgments.** This work has been partially supported by the Swiss National Science Foundation (SNF) under the grant number 200021\_143323/1.

## References

- [1] Noga Alon, Yossi Azar, Gerhard J. Woeginger, and Tal Yadid. Approximation schemes for scheduling. In *Proc. of the 8th Annual Symposium on Discrete Algorithms (SODA)*, pages 493–500, 1997.
- [2] Elliot Anshelevich, Anirban Dasgupta, Jon M. Kleinberg, Éva Tardos, Tom Wexler, and Tim Roughgarden. The price of stability for network design with fair cost allocation. *SIAM J. Comput.*, 38(4):1602–1623, 2008.
- [3] Arash Asadpour and Amin Saberi. On the Inefficiency Ratio of Stable Equilibria in Congestion Games. In *Proc. of the 5th International Workshop on Internet and Network Economics (WINE)*, volume 5929 of *LNCS*, pages 545–552, 2009.
- [4] Lawrence E. Blume. *Population games*. Addison-Wesley, 1998.
- [5] Ronald L. Graham. Bounds on multiprocessing timing anomalies. *SIAM Journal of Applied Mathematics*, 17(2):416–429, 1969.
- [6] Refael Hassin and Uri Yovel. Sequential scheduling on identical machines. *Oper. Res. Lett.*, 43(5):530–533, 2015.
- [7] Yasushi Kawase and Kazuhisa Makino. Nash equilibria with minimum potential in undirected broadcast games. *Theoretical Computer Science*, 482:33–47, 2013.
- [8] Elias Koutsoupias and Christos H. Papadimitriou. Worst-case equilibria. In *STACS 99, 16th Annual Symposium on Theoretical Aspects of Computer Science, Trier, Germany, March 4-6, 1999, Proceedings*, pages 404–413, 1999.
- [9] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, USA, 2007.